8

A Life-Size Example

	Page
Introduction	8-1
8A. The direct breakdown	8-1
8B. Looking at patterns	8-10
8C. We need not stop	8-15
8D. Looking still deeper	8-21
8E. And yet again	8-27
8F. Our analysis summarized	8-33
8G. Was the example gerrymandered?	8-40
8H. How far have we come?	8-40

Location of tables and plots for 8

Number	Page
1	8-2
2	8-4
3	8-5
4	8-7
5	8-8
6	8-11
7	8-12
8	8-14
9	8-16
10	8-18
11	8-19
12	8-20
13	8-22
14	8-24
15	8-25
16	8-26
17	8-28
18	8-29
19	8-31
20	8-32
21	8-35
22	8-36
23	8-38
24	8-39

8

A Life-Size Example

This chapter is devoted to a life-size example of what can and often does happen when we start to take a two-way table apart, using all the techniques we have so far set up.

To keep the arithmetic simple to do and easy to talk about, we use analysis by means and find our row-and-column description step by step. This is not done as a model -- given computing power it is usually much wiser to do analysis by midmean differences. It is done for easy thinking, and in the hope of having a somewhat different approach place what we have done before in a new light.

We introduce no new techniques in this chapter.

8A. The direct breakdown

Let us examine the % Democratic of the major party vote for President in 10 contiguous Nebraska counties in the 12 Presidential elections from 1920 to 1964 (data from Scammon 1965).

Table 1 sets out the numbers (and their sums, over these 10 counties) for each election. A moderately hard look at this table shows:

- substantial and seemingly rather consistent differences from election to election
- less notable differences, also seemingly consistent from county to county (with Howard the outstandingly different county).

table 1 of chapter 8

% of Major Party Vote for President Given to the Democratic Candidate --
10 Contiguous Nebraska Counties in 12 Elections

	Elections											
	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64
Adams	28.5	32.8	28.9	58.9	59.9	39.4	39.2	45.6	29.3	33.1	35.5	53.6
Buffalo	31.3	33.0	27.3	60.9	56.6	38.9	38.8	43.8	22.8	29.7	33.9	50.1
Dawson	29.9	33.6	25.1	61.2	52.9	34.0	31.2	38.2	20.3	23.9	26.9	45.3
Frontier	27.8	28.6	27.3	61.8	54.4	33.9	29.0	39.3	22.9	28.4	26.5	43.6
Hall	26.8	31.6	33.1	62.6	55.0	38.7	38.4	44.6	30.6	33.6	36.0	55.2
Hamilton	31.5	34.5	30.6	59.7	49.1	33.6	30.3	38.5	24.0	25.0	27.3	47.3
Howard	46.5	56.8	53.1	82.3	72.0	56.9	56.8	62.6	40.8	46.0	47.5	66.5
Merrick	31.1	32.9	30.0	62.9	50.4	33.7	34.1	37.9	22.7	28.3	31.0	49.2
Polk	34.1	34.3	32.5	64.2	52.8	38.1	39.2	40.8	23.3	31.4	33.1	51.8
Red Willow	37.3	36.8	33.2	63.8	62.4	44.3	40.7	46.2	26.9	29.9	33.1	46.9
Sums	324.8	354.9	321.1	638.3	565.5	391.5	377.7	437.5	263.6	309.3	330.8	509.5

Is there anything else notable here? To answer this question by placing one's nose (and eyes) a few inches above table 1 and moving it around, while looking and thinking hard, would not be easy. We need to do things to the numbers.

The obvious first step is to make allowance for election-to-election differences. The easiest thing to do is to find the arithmetic mean -- over the 10 counties -- for each election, and then to form residuals by subtracting the corresponding election mean from each observed %. The result is shown in table 2. Differences from county to county are now clearer, and it would be substantially easier to try to look beneath the surface to see what else may be going on. Again, a little arithmetic can clear the way appreciably. So let us use it.

Table 3 shows the effect of finding the mean -- over 12 elections -- of the first residuals for each county, and then subtracting the corresponding such mean from each first residual. We now have a mass of second residuals that are appreciably smaller than the numbers we have been so far examining. Nothing striking leaps to the eye, though there may well be further effects to be found.

As we have just now done, we shall find it convenient to call the residuals found by successive subtraction of pieces of a partial description "first residuals", "second residuals", "third residuals", and so on. We do this as a matter of convenience, not of principle -- indeed, what are "fourth residuals" by one chain of steps might be "seventh residuals"

table 2 of chapter 8

First Residuals (from election means) of % D for Table 1

(Election means, already subtracted)	'20 (32.5)	'24 (35.5)	'28 (32.1)	'32 (63.8)	'36 (56.6)	'40 (39.1)	'44 (37.8)	'48 (43.8)	'52 (26.4)	'56 (30.9)	'60 (33.1)	'64 (51.0)	(Sum) -
Adams	-4.0	-2.7	-3.2	-4.9	+3.3	+0.3	+1.4	+1.8	+2.9	2.2	2.4	2.6	(+2.1)
Buffalo	-1.2	-2.5	-4.8	-2.9	-0.0	-0.2	+1.0	-0.0	-3.6	-1.2	0.8	-0.9	(-15.5)
Dawson	-2.6	-1.9	-7.0	-2.6	-3.7	-5.1	-6.6	-5.6	-6.1	-7.0	-6.2	-5.7	(-60.6)
Frontier	-4.7	-6.9	-4.8	-2.0	-2.2	-5.2	-8.8	-4.5	-3.5	-2.5	-6.6	-7.4	(-58.1)
Hall	-5.7	-3.9	+1.0	-1.2	-1.6	-0.4	+0.6	+0.8	+4.2	+2.7	+2.1	+5.8	(+4.4)
Hamilton	-1.0	-1.0	-1.5	-4.1	-7.5	-5.5	-7.5	-5.3	-2.4	-5.9	-5.8	-3.7	(-51.0)
Howard	14.0	21.3	21.0	18.5	15.4	17.8	19.0	18.8	14.4	15.1	14.4	15.5	(205.4)
Merrick	-1.4	-2.6	-2.1	-0.9	-6.2	-5.4	-3.7	-5.9	-3.7	-3.6	-2.1	-1.8	(-39.4)
Polk	+1.6	-1.3	+0.4	+0.4	-3.8	-1.0	+1.4	-3.0	-3.1	0.5	0.0	0.8	(-7.2)
Red Willow	+4.8	+1.3	+1.1	-0.0	+5.8	+5.2	+2.9	+2.4	+0.5	-1.0	0.0	-4.1	(18.9)

Examples: 32.5 = 324.8/10, -4.0 = 28.5 - 32.5.

(Some computational errors can be found!)

table 3 of chapter 8

Second Residuals, After Subtraction of County Means

(Election means, already subtracted)	(County mean residual, already subtracted)	'20 (32.5)	'24 (35.5)	'28 (32.1)	'32 (53.8)	'36 (56.6)	'40 (39.1)	'44 (37.8)	'48 (43.8)	'52 (26.4)	'56 (30.9)	'60 (33.1)	'64 (51.0)
Adams	(+0.2)	-4.2	-2.9	-3.4	-5.1	+3.1	+0.1	1.2	1.6	2.7	2.0	2.2	2.4
Buffalo	(-1.3)	0.8	-1.2	-3.5	-1.6	1.3	1.1	2.3	1.3	-2.3	+0.1	2.1	0.4
Dawson	(-5.0)	+2.4	+3.1	-2.0	2.4	1.3	-0.1	-1.6	-0.6	-1.1	-2.0	-1.2	-0.7
Frontier	(-4.8)	+0.1	-2.1	+0.0	+2.8	2.6	-0.4	-4.0	0.3	1.3	2.3	-1.8	-2.6
Hall	(+0.4)	-6.1	-4.3	0.6	-1.6	-2.0	-0.8	0.2	0.4	3.8	2.3	1.7	5.4
Hamilton	(-4.3)	+3.3	+3.3	+2.8	+0.2	-3.2	-1.2	-3.2	-1.0	1.9	-1.6	-1.5	0.6
Howard	(17.1)	-3.1	+4.2	+3.9	+1.4	-1.7	+0.7	+1.9	+1.7	-2.7	-2.0	-2.7	-1.6
Merrick	(43.3)	1.9	0.7	1.2	2.4	-2.9	-2.1	-0.4	-2.6	-0.4	-0.3	1.2	1.5
Polk	(-0.6)	2.2	-0.7	1.0	1.0	-3.2	-0.4	2.0	-2.4	-2.5	1.1	+0.6	1.4
Red Willow	(+1.6)	3.2	-0.3	-0.5	-1.6	4.2	+3.6	+1.3	+0.8	-1.1	-2.6	-1.6	-5.7

Examples: +0.2 = 2.1/12; -1.3 = -15.6/12; -4.2 = -4.0 - (0.2); -1.2 = (-2.5) - (-1.3).

by another. We do it for one simple reason: because it gives us convenient, confusion-reducing labels for the various residuals that arise as we move along.

One natural thing to do at this stage is to look at the pattern of the values of these 120 second residuals, first as 120 numbers, and then with attention to how they are placed in the two-way table. The stem-and-leaf display of display 4 shows the pattern of values. Nothing very unusual is to be seen in the pattern as displayed, so we go on to find the hinges, the cutting points (which indicate needs for special attention) and the coding points (which determine the coded values). Display 5 shows the result. We thus find four of the second residuals beyond the cutting points and deserving special attention:

-6.1 for Hall county in 1920

-5.1 for Adams county in 1932

-5.7 for Red Willow county in 1964

+5.4 for Hall county in 1964

The piling up of the three largest of these four values in two counties and two elections -- the first and last of the 12 considered -- does nothing to reduce our feeling that they deserve attention.

Without detailed knowledge of Nebraska local politics, it is hard to follow this line much further. We ought to emphasize, however, that wisely chosen cutting points will direct special attention at a certain number of values that never prove to be abnormal. It is only by some sort of balance

display 4 of chapter 8

Stem-and-Leaf Display of the Distribution of Values of 120 Second Residuals from table 3

		Counts		
		Indiv.	Cumulative	
			Up	Down
-6	1	1	1	119
-5	17	2	3	117
-4	203	3	6	114
-3	452 212	6	12	108
-2	930 016 070 791 645 6	16	28	92
-1	266 128 620 657 661 6	16	44	76
-0	167 048 443 743 5	13	57	63
+0	181 413 624 267 768	15	72	48
+1	263 133 379 497 922 500 143	21	93	27
+2	702 431 448 633 842 0	16	109	11
+3	118 339 26	8	117	3
+4	22	2	119	1
+5	4	1	120	0
+6		0		

display 5 of chapter 8

Hinges, Cutting Points, and Coding Points
for the 120 Residuals

Hinges	Cutting Points	Coding Points	
			HIGH
		6.9	
			HIG
	5.2	5.2	
			·HI
		3.5	
			··H
+1.8		1.8	
			···
		0.1	
			···
-1.6		-1.6	
			L··
		-3.3	
			LO·
	-5.0	-5.0	
			LOW
		-6.7	
			LOWW

Hinges at ranks 30.5 and 90.5; difference = easy midspread. Cutting points one easy midspread further out. Coding points spaced one-half easy midspread.

between such "false positives" and their opposites -- the "false negatives" where truly exceptional values are not pin-pointed -- that we can do reasonably well with any automatic attention-calling process.

At this stage in the inquiry, then, we have the following clues to follow up:

- the typical behavior of this block of counties in each of the 12 elections (one number per election), which it would be natural to compare with the behavior of Nebraska as a whole and with the United States as a whole (as given, in particular, by election means)

- the typical behavior, in this block of elections, of each of these 10 counties (one number per county), which it would be natural to compare, one county with another (as given, in particular, by county mean residuals)

- the attention-deserving behavior of Hall county in 1920 and 1964, of Adams county in 1932, and of Red Willow county in 1964.

If our analysis were to stop here, we would certainly make these comparisons, and then ask all sorts of questions about the differences that appeared. Why, for one example, is Howard county so different (by being more than 15% more Democratic) than all the other nine counties? Does it depend less on farming? Does it do a different sort of farming? Are the national origins -- or religious associations -- of its people differently distributed? Or is there a long history of a -- once accidentally established -- strong Democratic

county organization? And so on, and on, not only for this difference but for many others.

The analysis we have so far reached of course corresponds to an elementary analysis by means. The corresponding elementary analysis is set out in table 6 as an explicit reminder.

8B. Looking at patterns

We now need to turn to the general aspects of the pattern of values of residuals in the two-way table. For this purpose, we need to summarize the values enough, but not too much. On one hand, the full detail of table 3 is much too detailed -- if there are patterns, they are likely to be hidden. On the other hand, the great summarization obtained by comparison with the two cutting points, thus putting every value into one of three classes, is much too condensed.

We need a convenient scheme of an intermediate degree of detail, such as the one provided by the coding points of display 5. Table 7 shows the results of applying this technique with both choices of columns: elections as columns and counties as columns.

The double appearance of the same information, the second time with rows and columns interchanged, is important. It is easy to compare LO and HI looking down a column, but hard to do a similar thing looking along a row.

At first glance, little may spring to the eye in this table, but if we look hard we will see quite a lot. Two

table 6 of chapter 8

The Second Residuals of table 3 in Elementary Analysis Form

	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	Election
Adams	-4.2	-2.9	-3.4	-5.1	3.1	0.1	1.2	1.6	2.7	2.0	2.2	2.4	+0.2
Buffalo	0.8	-1.2	-3.5	-1.6	1.3	1.1	2.3	1.3	-2.3	0.1	2.1	0.4	-1.3
Dawson	2.4	3.1	-2.0	2.4	1.3	-0.1	-1.6	-0.6	-1.1	-2.0	-1.2	-0.7	-5.0
Frontier	0.1	-2.1	0.0	2.8	2.6	-0.4	-4.0	0.3	1.3	2.3	-1.8	-2.6	-4.8
Hall	-6.1	-4.3	0.6	-1.6	-2.0	-0.8	0.2	0.4	3.8	2.3	1.7	5.4	+0.4
Hamilton	3.3	3.3	2.8	0.2	-3.2	-1.2	-3.2	-1.0	1.9	-1.6	-1.5	0.6	-4.3
Howard	-3.1	4.2	3.9	1.4	-1.7	0.7	1.9	1.7	-2.7	-2.0	-2.7	-1.6	+17.1
Merrick	1.9	0.7	1.2	2.4	-2.9	-2.1	-0.4	-2.6	-0.4	-0.3	1.2	1.5	-3.3
Polk	2.2	-0.7	1.0	1.0	-3.2	-0.4	2.0	-2.4	-2.5	1.1	+0.6	1.4	-0.6
Red Willow	3.2	-0.3	-0.5	-1.6	4.2	3.6	1.3	0.8	-1.1	-2.6	-1.6	-5.7	+1.6
County	-6.9	-3.9	-7.3	+14.4	+17.2	-0.3	-1.6	+4.4	-13.0	-8.5	-6.3	11.6	39.4

table 7 of chapter 8

Simply Coded Representation of the Second Residuals of table 3

	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	Elect
Adams	LO·	L··	L··	LOW	··H	·····	···	···	··H	··H	··H	··H	+0.2
Buffalo	··H	···	LO·	···	···	···	··H	···	L··	·····	··H	···	-1.3
Dawson	··H	··H	L··	··H	···	···	···	···	···	L··	···	···	-5.0
Frontier	·····	L··	···	··H	··H	···	LO·	···	···	··H	L··	L··	-4.8
Hall	LOW	LO·	···	···	L··	···	···	···	·HI	··H	··H	HIG	+0.4
Hamilton	··H	··H	··H	···	L··	···	L··	···	··H	···	···	···	-4.3
Howard	L··	·HI	·HI	···	···	···	··H	···	L··	L··	L··	···	17.1
Merrick	··H	···	···	··H	L··	L··	···	L··	···	···	···	···	-3.3
Polk	··H	···	···	···	L··	···	··H	L··	L··	···	···	···	-0.6
Red W'lw	··H	···	···	···	·HI	·HI	···	···	···	L··	···	LOW	+1.6
County	-6.9	-3.9	-7.3	14.4	17.2	-0.3	-1.6	4.4	-13.0	-8.5	-6.3	11.6	39.4

	Adams	Buffalo	Dawson	Frontier	Hall	Hamilton	Howard	Merrick	Polk	Red W'lw	County
'20	LO·	··H	··H	·····	LOW	··H	L··	··H	··H	··H	-6.9
'24	L··	···	··H	L··	LO·	··H	·HI	···	···	···	-3.9
'28	L··	LO·	L··	···	···	··H	·HI	···	···	···	-7.3
'32	LOW	···	··H	··H	···	···	···	··H	···	···	14.4
'36	··H	···	···	··H	L··	L··	···	L··	L··	·HI	17.2
'40	·····	···	···	···	···	···	···	L··	···	·HI	-0.3
'44	·····	··H	···	LO·	···	L··	··H	···	··H	···	-1.6
'48	·····	···	···	···	···	···	···	L··	L··	···	+4.4
'52	··H	L··	···	···	·HI	··H	L··	···	L··	···	-13.0
'56	··H	·····	L··	··H	··H	···	L··	···	···	L··	-8.5
'60	··H	··H	···	L··	··H	···	L··	···	···	···	-6.3
'64	··H	···	···	L··	HIG	···	···	···	···	LOW	11.6
Elec.	+0.2	-1.3	-5.0	-4.8	+0.4	-4.3	17.1	-3.3	-0.6	+1.6	39.4

counties, Adams and Hall, seem to show a trend with time away from the Republicans and toward the Democrats. Conversely, Howard, and probably Red Willow, show an opposite trend.

These trends account in a simple way for the three attention-deserving values we noted at the end of the previous section (8A) and will replace them in our mental summary.

Once we start to make such quasi-pictorial, quasi-numerical tables, however, we are driven to go further -- to ask ourselves why the rows should remain in the order in which they have first appeared. (There is some excuse for the order of columns, which is the order of increasing date.) It seems most unlikely that alphabetic order of county name reflects a useful political reality. Rearranging these counties badly needs trying. While it might be better to arrange them from east to west, from north to south, or from wet to dry, we are going to begin with an even simpler arrangement: from Republican to Democratic, using the county mean residuals or county adjustments as the basis of rearrangement.

Having rearranged rows it is natural to consider rearranging columns. Again, many orders could be used -- the initial, chronological order must clearly be looked at whenever relevant -- but the one that requires no outside information is to arrange columns, also, from Republican to Democratic. Table 8 shows the result. In this instance, rearrangement has not done too much to alter the picture.

If we were to stop our analysis here, we might summarize our clues as follows:

table 8 of chapter 8

Rearranged Simply Coded Representations of the Second Residuals of table 3

39.4	-13.0	-8.5	-7.3	-6.9	-6.3	-3.9	-1.6	-0.3	+4.4	11.6	14.4	17.2	County
17.1	L··	L··	·HI	L··	L··	·HI	··H	···	···	L··	···	···	Howard
1.6	···	L··	···	··H	···	···	···	·HI	···	···	LOW	···	Red W'lw
0.4	·HI	··H	···	LOW	··H	LO·	···	···	···	···	HIG	···	Hall
0.2	··H	··H	L··	LO·	··H	L··	···	·····	···	···	··H	LOW	Adams
-0.6	L··	···	···	··H	···	···	··H	···	L··	···	···	···	Polk
-1.3	L··	·····	LO·	···	··H	···	··H	···	···	···	···	···	Buffalo
-3.3	L··	···	···	··H	···	···	···	L··	L··	L··	···	··H	Merrick
-4.3	·HI	···	··H	··H	···	··H	L··	···	···	···	···	···	Hamilton
-4.8	···	··H	···	·····	L··	··H	LO·	···	···	···	···	··H	Frontier
-5.0	···	L··	L··	··H	···	··H	···	···	···	···	···	··H	Dawson
Elect	'52	'56	'28	'20	'60	'24	'44	'40	'48	'64	'32	'36	

39.4	-5.0	-4.8	-4.3	-3.3	-1.3	-0.6	0.2	0.4	1.6	17.1	Election
17.2	···	···	···	LOW	···	··H	··H	···	··H	··H	'36
14.4	···	LOW	HIG	··H	···	···	···	···	···	···	'32
11.6	L··	···	···	···	···	···	···	···	···	···	'64
4.4	···	···	···	···	L··	···	L··	···	···	···	'48
-0.3	···	·HI	···	·····	···	···	L··	···	···	···	'40
-1.6	··H	···	···	···	··H	··H	···	L··	LO·	···	'44
-3.9	·HI	···	LO·	L··	···	···	···	··H	··H	··H	'24
-6.3	L··	···	··H	··H	···	··H	···	···	···	···	'60
-6.9	L··	··H	LOW	LO·	··H	···	··H	··H	·····	··H	'20
-7.3	·HI	···	···	L··	···	LO·	···	··H	···	L··	'28
-8.5	L··	L··	··H	··H	···	·····	···	···	··H	L··	'56
-13.0	L··	···	·HI	··H	···	L··	L··	·HI	···	···	'52
County	Dawson	Fron'r	Ham'n	Merr'k	Buff'o	Polk	Adams	Hall	Red W'w	Howard	

- the typical behavior of these counties in 12 elections

- the typical behavior in these elections of 10 counties

- time trends of notable size in some counties, toward the Democrats in Adams and Hall, toward the Republicans in Red Willow, and, probably, Howard.

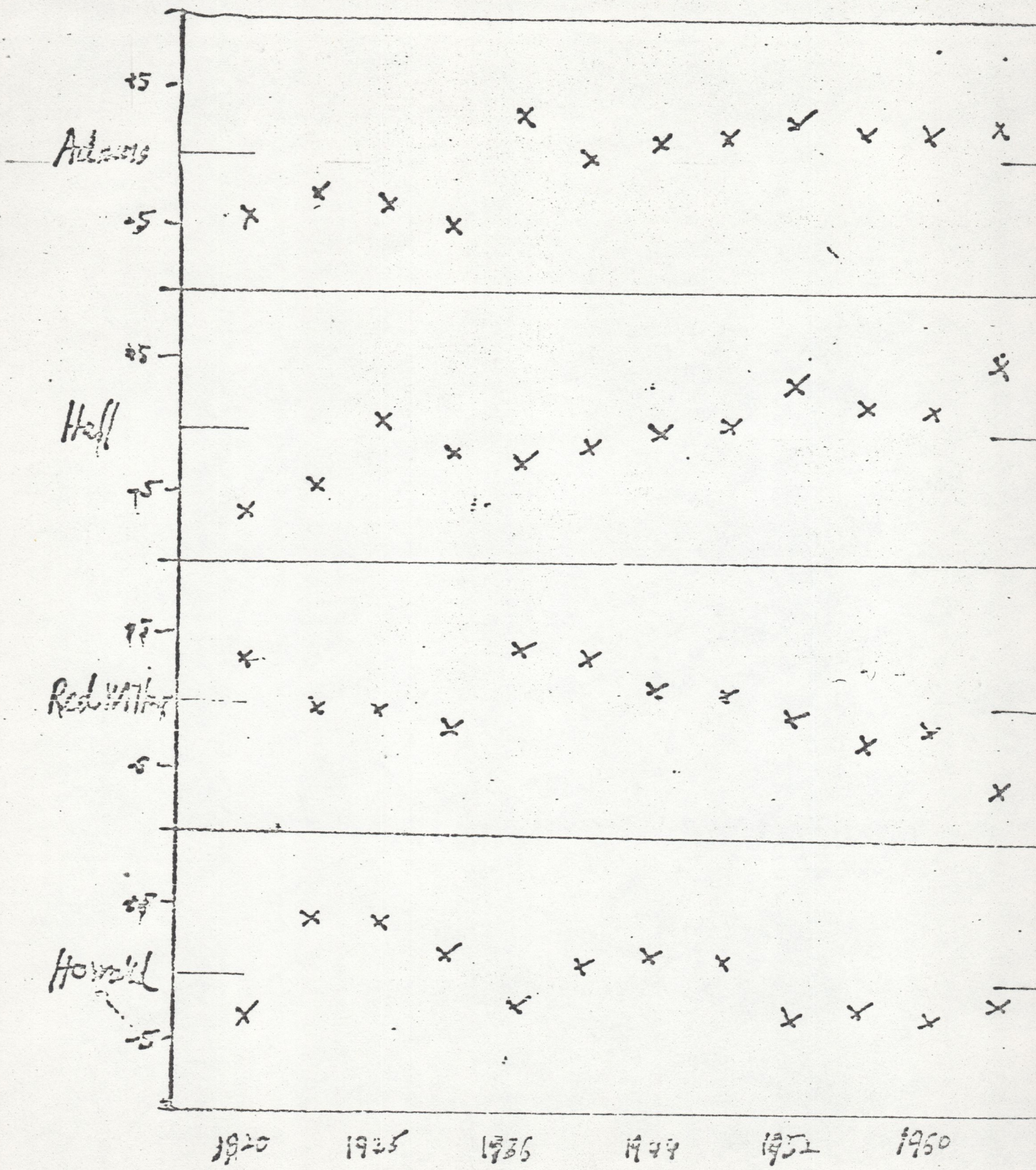
If we were to stop here, we would make all the comparisons mentioned at the close of section 8A, ask all the questions about typical behavior that would have been asked at that point, and go on and ask many political questions about time trends.

8C. <u>We need not stop</u>

In this example, the most prominent effects, once we clear away rows (= counties) and columns (= elections) were the apparent time trends in Adams, Hall, Red Willow, and, possibly, Howard counties. We saw they seemed to be there, but we have not looked AT them. Plot 9, which shows (second) residuals against date for each of these four counties, remedies this failure. When we look at these four plots it is quite clear that there were time trends. Why were they there? We ought at least to ask what we can learn easily about changes in those counties over these years.

Before we turn to ask economic, social, ethnic, or religious questions about the differences between these counties, we ought to consider grasping at any straws our source of voting data offers us. In addition to the % split of the

Behavior of residuals over time for 4 counties



vote, this source offers us information on total vote. While differences in party activity can affect turnout in individual elections, long-time trends in total vote probably indicate long-term changes in county populations. Accordingly it seems sensible to inquire whether differences in % D trends are associated with differences in total vote trends. Table **10** gives the figures, both for individual elections and for mid-means of early and late thirds. Clearly the rates of growth of votes -- which presumably reflect the rates of growth of population -- are quite different for the four counties: about 32% for Adams and about 60% for Hall, as against about 17% for Red Willow and about -17% for Howard. The vote in the two counties with noticeably increasing residuals grew notably faster than the vote in the counties with noticeably decreasing residuals. Accordingly, we ought to look further.

Table **11** presents the midmeans for early and late thirds of the 12 elections for both residuals and total votes. Plot **12** plots the change in midmean residual against the % change in midmean total vote. We see a definite appearance of increasing (= more Democratic) residuals with increasing growth in total vote. The county of Dawson, marked D to the lower right, might be enough different in behavior to deserve special consideration.

We can now add many questions about the economic, social, ethnic and religious factors which may carry the effects of different rates of population growth into different rates of trend in % Democratic.

table 10 of chapter 8

Total Presidential Election Vote in Four Counties

	Adams	Hall	Red Willow	Howard
1920	7,008	7,123	3,385	2,971
1924	8,516	8,525	4,329	3,455
1928	10,186	10,306	5,384	4,157
1932	9,934	10,492	5,592	4,306
1936	10,621	11,810	5,692	4,494
1940	10,941	12,099	5,600	3,934
1944	11,777	12,419	5,239	3,598
1948	10,212	10,284	4,848	3,026
1952	12,778	15,043	6,064	3,571
1956	12,234	14,351	5,431	3,150
1960	12,296	15,259	5,812	3,194
1964	12,027	14,988	5,156	3,044
(midmean 1920 to 1932)	(9,225)	(9,416)	(4,856)	(3,806)
(midmean 1952 to 1964)	(12,265)	(15,015)	(5,622)	(3,172)

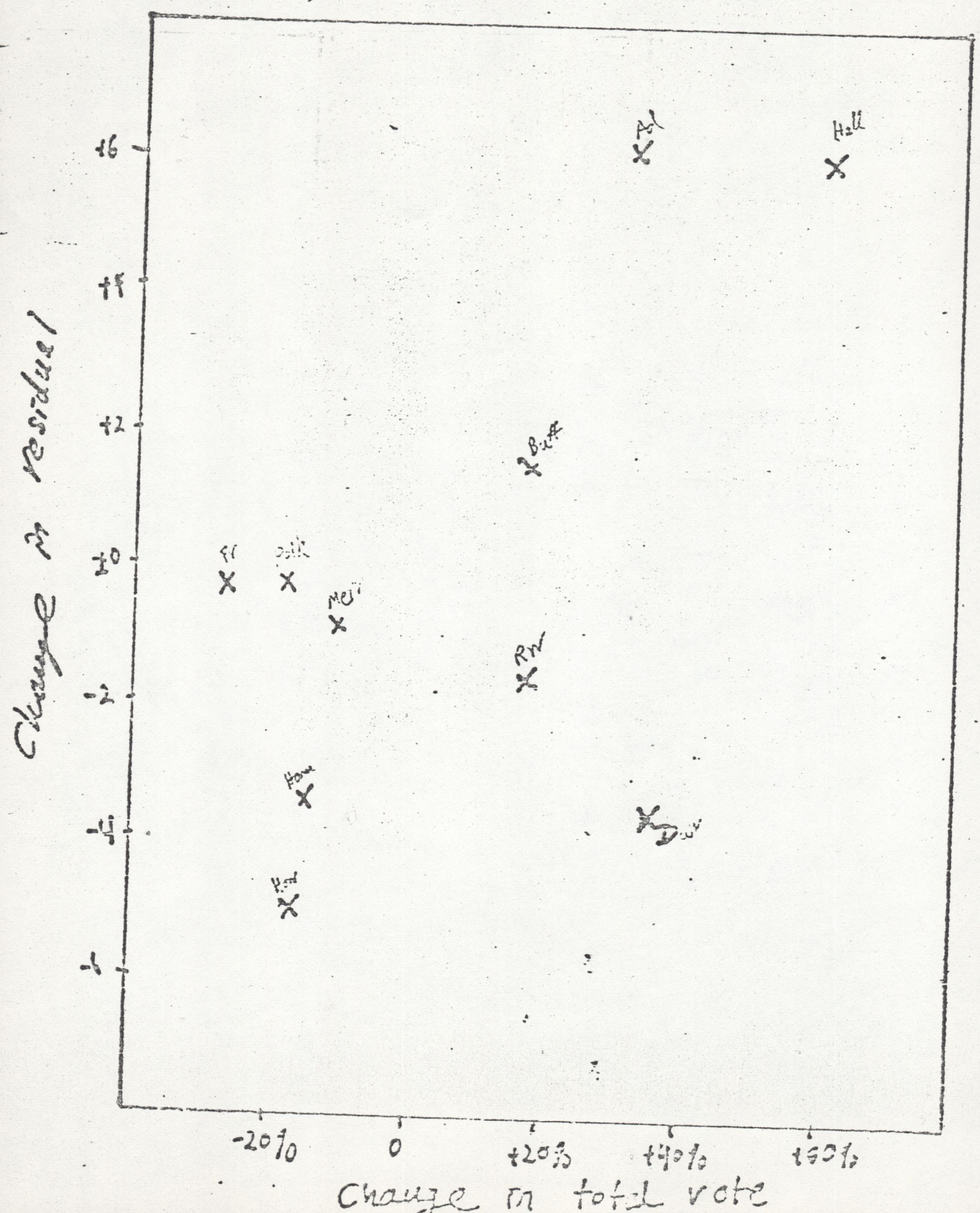
table 11 of chapter 8

Midmeans for Early (1920-1932) and Late (1952-1964) Thirds of the Set of 12 Presidential Elections for Both the (Second) Residuals of Table 3 and the Total Vote, With Changes from Early to Late

County	Midmean residuals			Midmean total votes		
	Early	Late	Diff	Early	Late	% Diff
Adams	-3.0	2.3	+5.3	9,225	12,265	+32%
Buffalo	-1.4	0.2	+1.6	9,347	10,914	+17%
Dawson	2.4	-1.2	-3.6	6,427	8,705	+35%
Frontier	0.0	-0.2	-0.2	3,059	2,216	-27%
Hall	-3.0	3.0	+6.0	9,416	15,015	+60%
Hamilton	3.0	-0.4	-3.4	5,140	4,382	-15%
Howard	2.6	-2.4	-5.0	3,806	3,172	-17%
Merrick	1.6	0.4	-1.2	4,486	3,980	-11%
Polk	1.0	0.8	-0.2	4,383	3,599	-18%
Red Willow	-0.4	-2.1	-1.7	4,856	5,622	+14%

plot 12 of chapter 8

Comparison of changes from 1920-1932 to 1952-1964 for the 10 counties



8D. Looking still deeper

To look still deeper, we have again to "clear the trail". We have recognized the existence of appreciable time trends; we must admit that their presence could conceal other phenomena. Are there other phenomena? To see the answer to such a question we need to strip away whatever we have recognized -- here, visible time trends. So let us begin.

To calculate how much time trend appears to be present -- if this trend follows a straight-line course -- is easy. If the trend were

-11,-9,-7,-5,-3,-1,+1,+3,+5,+7,+9,+11

which is surely straight-line and arithmetically simple, the midmeans of the early and late thirds would be -8 and +8, respectively, so that their difference would be +16.

For Buffalo county (see table 11) the observed difference of the corresponding midmeans is +1.6. Clearly we may adjust Buffalo county by subtracting one-tenth of the values (-11 in 1920 to +11 in 1964) we took as an example. Clearly we should act proportionately in other counties. In Howard county, for instance, we should subtract - (5.0)/16 times the simple values (that is, add 5/16 times them).

Table 13 sets out the resulting adjusted (third) residuals, which appear somewhat smaller, overall, than the unadjusted residuals. Indeed, the two sets of hinges and easy midspread are:

table 13 of chapter 8

Residuals After Removing Linear Time Trends From Residuals of Table 3
(Third residuals)

County \ Election	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64
Adams	0.0	0.5	-0.7	-3.2	4.2	0.5	0.7	0.5	0.8	-0.7	-1.2	-1.8
Buffalo	1.9	-0.3	-2.8	-1.1	1.6	1.2	2.2	1.0	-2.7	-0.6	1.2	-0.7
Dawson	-0.1	1.1	-3.6	1.3	0.6	-0.3	-1.4	0.1	0.0	-0.4	+0.8	+1.8
Frontier	0.0	-2.2	-0.1	2.7	2.6	-0.4	-4.0	0.3	1.4	2.4	-1.7	-2.5
Hall	-2.0	-0.9	-3.2	-3.5	-0.9	-0.4	-0.2	-0.7	1.9	-0.3	-1.7	1.3
Hamilton	0.9	1.4	1.3	-0.9	-3.8	-1.5	-3.0	-0.4	3.0	-0.1	0.4	2.9
Howard	-6.5	1.4	1.7	-0.2	-2.6	0.4	2.2	2.6	-1.1	0.2	0.1	1.8
Merrick	1.1	0.0	0.7	2.0	-3.1	-2.2	-0.3	-2.4	0.0	0.2	1.9	2.3
Polk	2.1	-0.8	0.9	0.9	-3.2	-0.4	2.0	-2.4	-2.4	1.2	0.7	1.5
Red W"lw	2.0	-1.3	-1.2	-2.1	3.9	3.5	1.4	1.1	-0.7	-1.9	-0.6	-4.5

Coding points:

	LOWW		L··			··H
-4.95		-1.2		2.55		
	LOW		···			·HI
-3.7		0.05		3.8		
	LO·		···			HIG
-2.45		+1.3		4.05		
	L··		··H			HIGH

	L hinge	U hinge	Midspread
Residuals	-1.6	1.8	3.4
Detrended residuals	-1.2	1.3	2.5

so that we have made the typical residual only about 5/7 of its previous size by taking out linear time trends. (As a result the square of a typical residual is only about half, 25/49, of its former size.)

To get a clearer idea of what these residuals are trying to say, we need to code them in the usual way. Skipping any table with alphabetically arranged counties, tables 14 and 15 show the coded residuals with counties in Republican to Democratic order and with elections once in order of date and once in Republican to Democratic order. Looking along rows and down columns shows no particular trends. Have we come to the end of the road?

We do see something else, however: a tendency for the bunching up of large residuals in few elections. Clearly 1936, and probably 1964, have an unfair share of large residuals. We should, therefore, try to make sense out of their pattern. What could we try to relate them to? Probably not any voting information. What else do we know but have failed to use?

The simplest answer to this question is "geography". Map 16 shows the coded residuals for 1936 and 1964, which clearly show geographical coherence. Whatever happened in those elections that we have not yet described, happened over areas larger than counties.

table 14 of chapter 8

Coded Residuals from table 13 (Chronological Order of Elections)

Elect	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	
-5.0	···	···	LO·	···	···	···	L··	···	···	···	···	··H	Dawson
-4.8	···	L··	···	·HI	·HI	···	LOW	···	··H	··H	L··	L··	Frontier
-4.3	···	··H	···	···	LOW	L··	LO·	···	·HI	···	···	·HI	Hamilton
-3.3	···	···	···	··H	LO·	L··	···	L··	···	···	··H	··H	Merrick
-1.3	··H	···	LO·	···	··H	···	··H	···	LO·	···	···	···	Buffalo
-0.6	··H	···	···	···	LO·	···	··H	L··	L··	···	··H	··H	Polk
0.2	···	···	···	LO·	HIGH	···	···	···	···	···	···	L··	Adams
0.4	L··	···	LO·	LO·	···	···	···	···	··H	···	L··	···	Hall
1.6	··H	L··	···	L··	HIG	·HI	··H	···	···	L··	···	LOW	Red Willow
17.1	LOWW	··H	··H	···	LO·	···	··H	·HI	···	···	···	··H	Howard
39.4	-6.9	-3.9	-7.3	14.4	17.2	-0.3	-1.6	4.4	-13.0	-8.5	-6.3	11.6	County

Elect	-5.0	-4.8	-4.3	-3.3	-1.3	-0.6	0.2	0.4	1.6	17.1	39.4
'20	···	···	···	···	··H	··H	···	L··	··H	LOWW	-6.9
'24	···	L··	··H	···	···	···	···	···	L··	··H	-3.9
'28	LO·	···	···	···	LO·	···	···	LO·	···	··H	-7.3
'32	···	·HI	···	··H	···	···	LO·	LO·	L··	···	14.4
'36	···	·HI	LOW	LO·	··H	LO·	HIGH	···	HIG	LO·	17.2
'40	···	···	L··	L··	···	···	···	···	·HI	···	-0.3
'44	L··	LOW	LO·	···	··H	··H	···	···	··H	··H	-1.6
'48	···	···	···	L··	···	L··	···	···	···	·HI	4.4
'52	···	··H	·HI	···	LO·	L··	···	··H	···	···	-13.0
'56	···	··H	···	···	···	···	···	···	L··	···	-8.5
'60	···	L··	···	··H	···	··H	···	L··	···	···	-6.3
'64	··H	L··	··HI	··H	···	··H	L··	···	LOW	··H	11.6
	Dawson	Fron'r	Ham'n	Merr'k	Buff'o	Polk	Adams	Hall	Red W'w	Howard	County

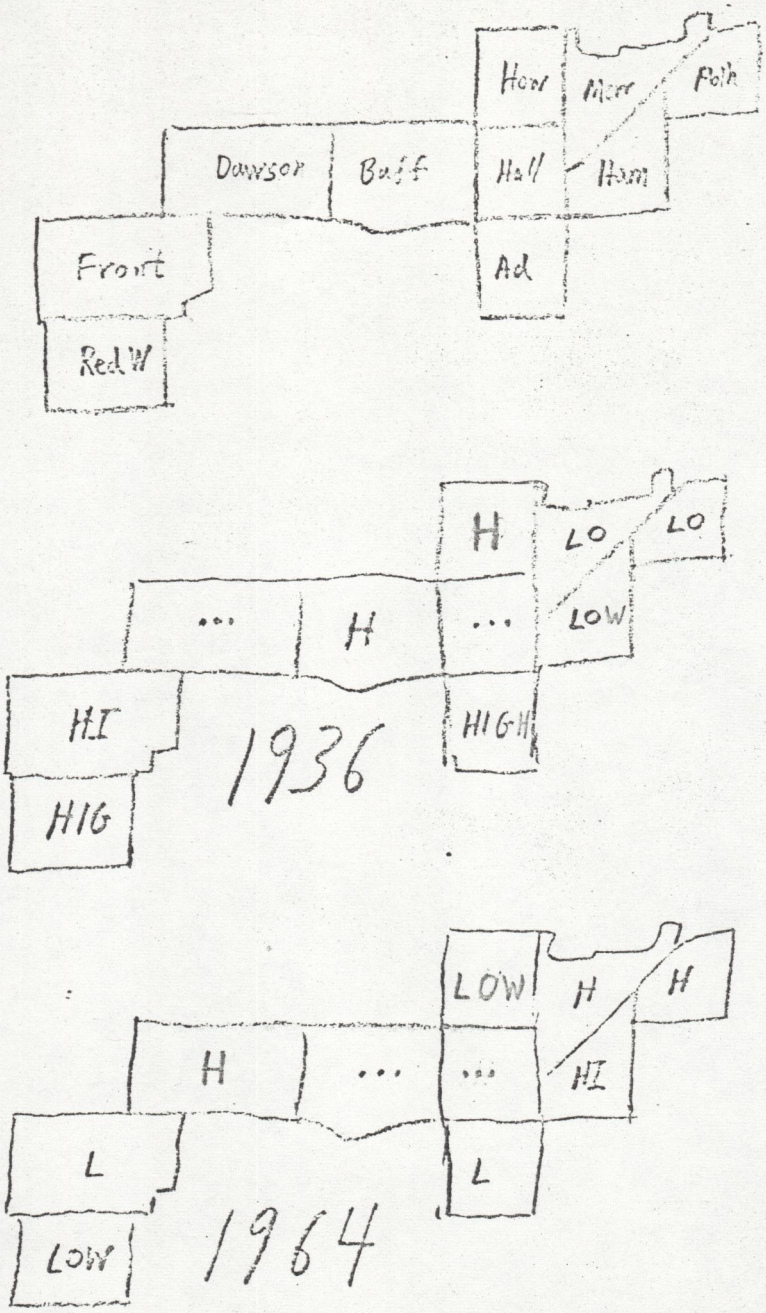
table 15 of chapter 8

Coded Residuals from table 13 (Elections and Counties from Rep to Dem)

39.4	-13.0	-8.5	-7.3	-6.9	-6.3	-3.9	-1.6	-0.3	4.4	11.6	14.4	17.2	County
-5.0	···	···	LO·	···	···	···	L··	···	···	··H	···	···	Dawson
-4.8	··H	··H	···	···	L··	L··	LOW	···	···	L··	·HI	·HI	Frontier
-4.3	·HI	···	···	···	···	··H	LO·	L··	···	·HI	···	LOW	Hamilton
-3.3	···	···	···	···	··H	···	···	L··	L··	··H	··H	LO·	Merrick
-1.3	LO·	···	LO·	··H	···	···	··H	···	···	···	···	··H	Buffalo
-0.6	L··	···	···	··H	··H	···	··H	···	L··	··H	···	LO·	Polk
0.2	···	···	···	···	···	···	···	···	···	L··	LO·	HIGH	Adams
0.4	··H	···	LO·	L··	L··	···	···	···	···	···	LO·	···	Hall
1.6	···	L··	···	··H	···	L··	··H	·HI	···	LOW	L··	HIG	Red Willow
17.1	···	···	··H	LOWW	···	··H	··H	···	·HI	··H	···	LO·	Howard
Elect	'52	'56	'28	'20	'60	'24	'44	'40	'48	'64	'32	'36	

39.4	-5.0	-4.8	-4.3	-3.3	-1.3	-0.6	0.2	0.4	1.6	17.1	Elections
-13.0	···	··H	·HI	···	LO·	L··	···	··H	···	···	'52
-8.5	···	··H	···	···	···	···	···	···	L··	···	'56
-7.3	LO·	···	···	···	LO·	···	···	LO·	···	··H	'28
-6.9	···	···	···	···	··H	··H	···	L··	··H	LOWW	'20
-6.3	···	L··	···	··H	···	··H	···	L··	···	···	'60
-3.9	···	L··	··H	···	···	···	···	···	L··	··H	'24
-1.6	L··	LOW	LO·	···	··H	··H	···	···	··H	··H	'44
-0.3	···	···	L··	L··	···	···	···	···	·HI	···	'40
4.4	···	···	···	L··	···	L··	···	···	···	·HI	'48
11.6	··H	L··	·HI	··H	···	··H	L··	···	LOW	··H	'64
14.4	···	·HI	···	··H	···	···	LO·	LO·	L··	···	'32
17.2	···	·HI	LOW	LO·	··H	LO·	HIGH	···	HIG	LO·	'36
Counties	Dawson	Fron'r	Ham'n	Merr'k	Buff'o	Polk	Adams	Hall	Red W'w	Howard	

Abbreviated County names, and behavior of coded third residuals in 1936 and 1964



(Coded residuals from Table 14 or Table 15)

8E. And yet again

At the end of the last section we found something new associated with counties. In a two-way table, something new about one way almost always offers a new way to look at what happens along the other way, a new handle for further inquiry. In this case, something new about counties offers us a chance to look for something new about elections. Let us grasp this handle, and see what it can do for us.

We found that 1936 and 1964 show striking geographical patterns almost opposite in character. If we isolate a few counties whose behavior is near one extreme and compare them with a few counties near the other extreme, we can have a new measure of elections. (One that is sure to have 1936 and 1964 close to opposite extremes.)

We could select these counties in terms of their behavior in 1936, or their behavior in 1964, but we may do still better if we combine these two elections by looking at the differences between them. The 1964 minus 1936 differences in % D for the detrended (= third) residuals for the 10 counties are: -8.4,-6.0,-5.1,-2.4,-2.3,+2.2,+4.4,+4.7,+5.4,+6.7. If we were to seek out a tilt for each individual election, we ought to take the counties with -8.4,-6.0,-5.1 for the lower portion and those with +4.4,+4.7,+5.4 and +6.7 for the upper. Table 17 shows the calculation.

The numbers seem interesting enough to be worth looking at. Accordingly plot 18 shows their behavior as a function of time. The picture is surprisingly clear:

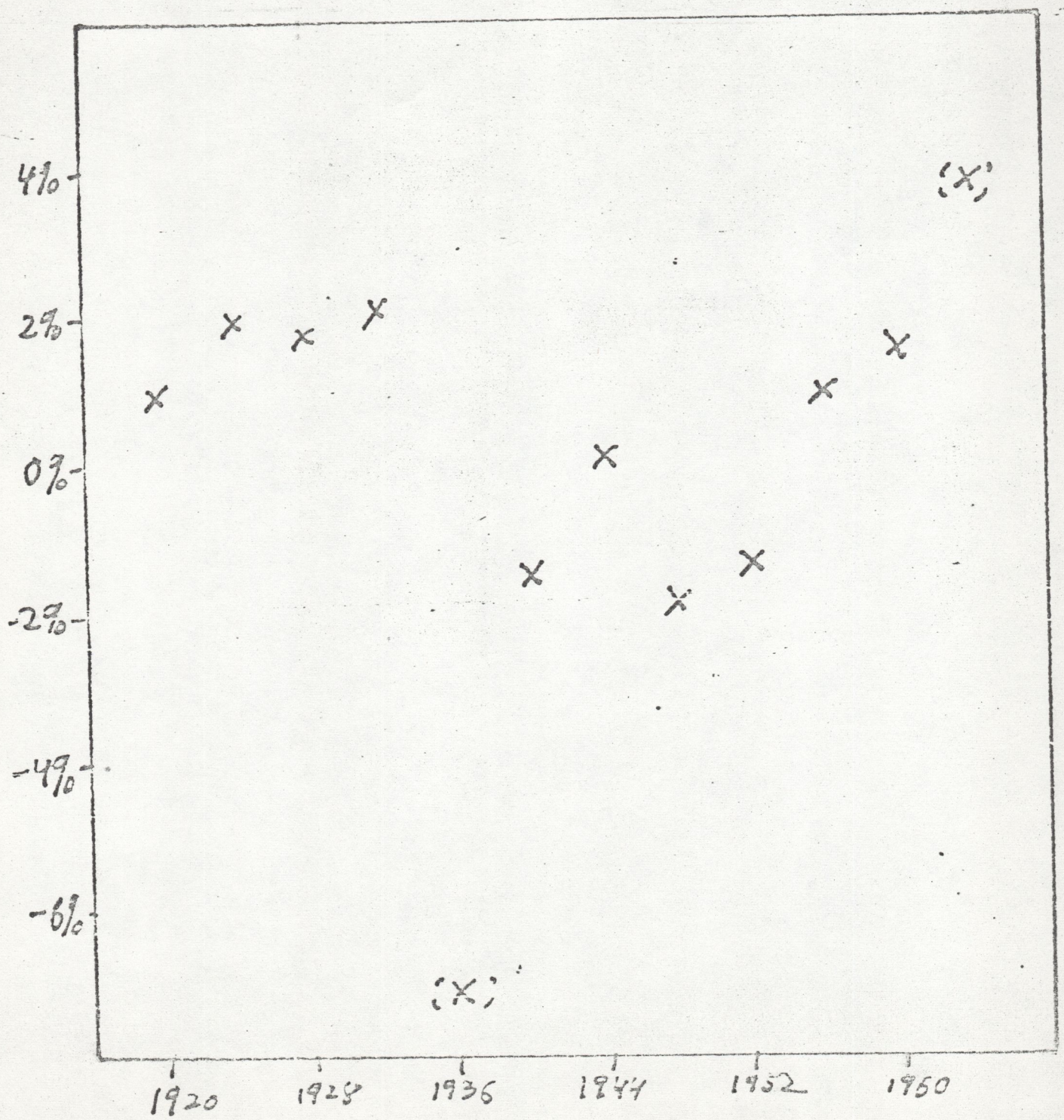
table 17 of chapter 8

Assessment of Relative Tilt for Each Election with Reference to the 1964-1936 Difference

	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64
-8.4 Red Willow	2.0	-1.3	-1.2	-2.1	3.9	3.5	1.4	1.1	-0.7	-1.9	-0.6	-4.5
-6.0 Adams	0.0	0.5	-0.7	-3.2	4.2	0.5	0.7	0.5	0.8	-0.7	-1.2	-1.8
-5.1 Frontier	0.0	-2.2	-0.1	2.7	2.6	-0.4	-4.0	0.3	1.4	2.4	-1.7	-2.5
(midmean)	(0.0)	(-1.3)	(-0.7)	(-2.1)	(3.9)	(0.5)	(0.7)	(0.5)	(0.8)	(-0.7)	(-1.2)	(-2.5)
+4.4 Howard	-6.5	1.4	1.7	-0.2	-2.6	0.4	2.2	2.6	-1.1	+0.2	+0.1	+1.8
+4.7 Polk	2.1	-0.8	0.9	0.9	-3.2	-0.4	2.0	-2.4	-2.4	1.2	0.7	1.5
+5.4 Merrick	1.1	0.0	0.7	2.0	-3.1	-2.2	-0.3	-2.4	0.0	0.2	1.9	2.3
+6.7 Hamilton	0.9	1.4	1.3	-0.9	-3.8	-1.5	-3.0	-0.4	3.0	-0.1	+0.4	2.9
(midmean)	(1.0)	(0.7)	(1.1)	(0.4)	(-3.2)	(-1.0)	(0.8)	(-1.4)	(-0.6)	(0.2)	(0.6)	(2.0)
Diff of midmeans	1.0	2.0	1.8	2.5	-7.1	-1.5	0.1	-1.9	-1.4	0.9	1.8	4.5

plot 18 of chapter 8

Difference of midyears of detrended residuals
4 Northerly countries (1964 - 1936 >> 0)
referred to
3 Southerly countries (1964 - 1936 << 0)



- the first four elections (1920-1932) show a rather stable value.

- 1936 snaps violently in the opposite direction (it will be unfairly low, since it helped determine the extreme sets of counties used as the portions).

- the next six elections (1940-1960) seem to be relatively consistent, though there may be an upward trend. Their level is appreciably lower than that of the first four elections.

- 1964 definitely either jumps up or continues an upward trend (it, too, will be unfairly extreme because of its use in choosing the portions).

We have found another clue. It seems likely we are at the bottom of the bag. But we may as well take another look. We have just identified two sets of relatively peaceful elections -- 1920-1932 and, say 1940-1952 -- as differing in behavior. How does this appear in terms of individual counties? Table 19 shows the calculation, and map 20 the geographical pattern. So far as these elections are concerned, there seems to be a line between the three Eastern counties and the other seven.

We need to stop and give ourselves a word of caution now. When we fitted time trends to our (second) residuals (see table 11) we compared 1920-1932 with 1952-1964, choosing these as early and late. After making that adjustment, we have compared 1920-1932 with 1940-1952 because these seemed to be quiet times. Surely there must be some overlap, some

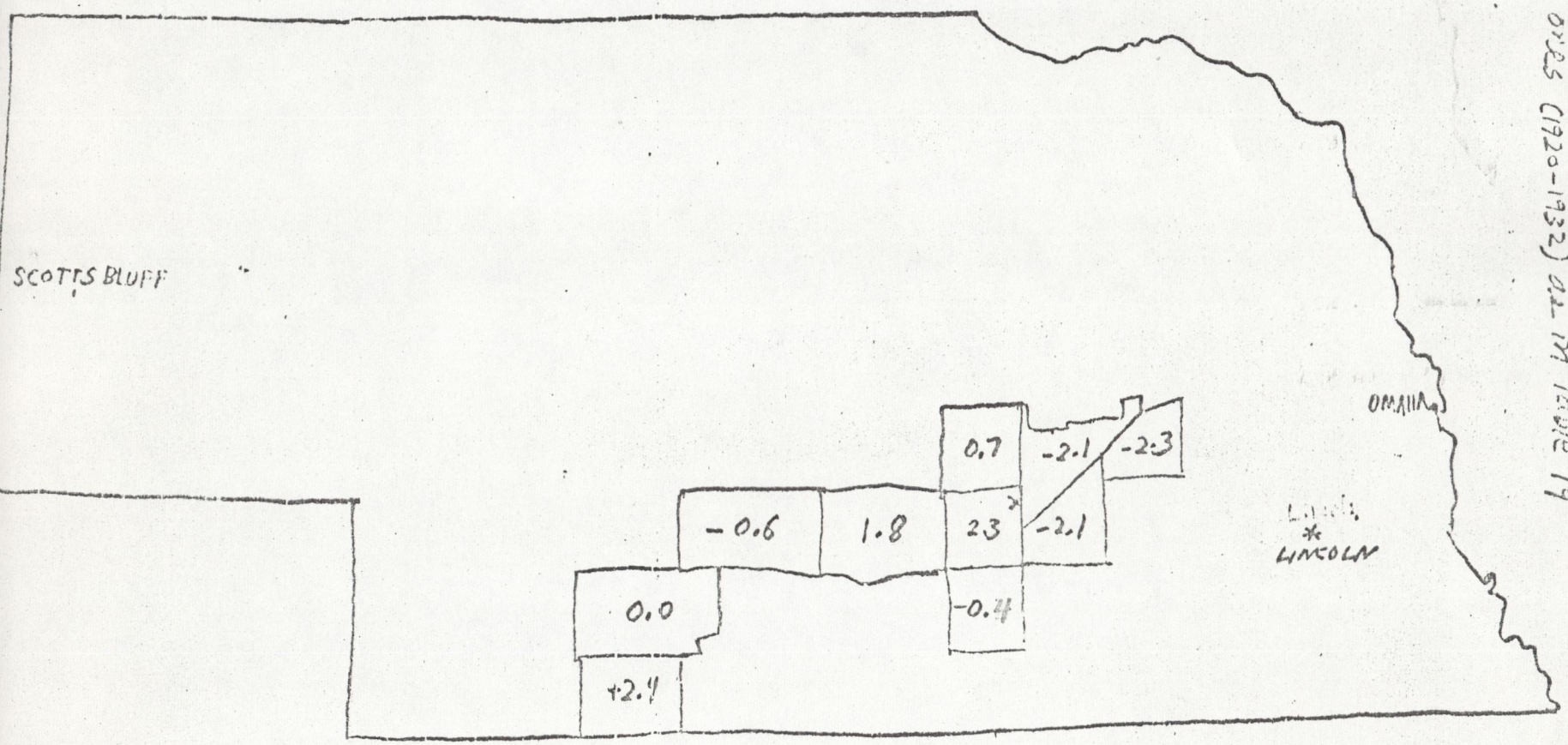
table 19 of chapter 8

Comparison of Early Quiet and Late Quiet Elections

County	Midmean 1920-1932	Midmean 1940-1952	Difference
Adams	-0.4	0.6	1.0
Buffalo	-0.7	1.1	1.8
Dawson	0.5	-0.1	-0.6
Frontier	0.0	0.0	0.0
Hall	-2.6	-0.3	2.3
Hamilton	1.1	-1.0	-2.1
Howard	0.6	1.3	0.7
Merrick	0.9	-1.2	-2.1
Polk	0.9	-1.4	-2.3
Red Willow	-1.2	1.2	2.4

map 20 of chapter 8

Location of counties in Nebraska and results of comparing later quiet elections (1940-1952) with earlier ones (1920-1932) as in Table 19



x is location of Grand Island (in Hall county)

tendency for the first fitting to try to make part of the adjustment that ought to be made by the second one. This is a reasonable warning -- we all ought to heed it by caution in making interpretations, but we shall do nothing more about it here.

There is another step we could take. This is to clear our trail in another way by eliminating the two most unusual elections, 1964 and 1936, from all consideration. This would leave us with a 10x10 table, which could then be analyzed in the way we have treated the 10x12 table. The changes in residuals would, however, be small, only up to ±0.2. We leave this further explorative step to the interested reader.

8F. Our analysis summarized

It is now time to try to put our results into the most suggestive forms we can. After all, they are but clues, intended to raise political questions for subject-matter inquiry. They will serve their functions as clues most effectively if they are presented as suggestively as possible.

What do we need to present, and where is the raw material?

- Ten-county behavior, election by election (table 1 or table 2)
- Northeasterly four counties minus southwesterly three counties, election by election (table 17, plot 18)
- Twelve-election behavior, county by county (table 2 or 3)
- Time trends, county by county (table 11)

- 1964 detrended residual minus 1936 detrended residual, county by county (from table 13)

- 1940-52 midmean detrended residual minus 1920-32 midmean detrended residual (table 19)

- Relation of time trends to vote growth (plot 12)

- The unusual behavior of Howard county in 1920.

Where one can, the way to be suggestive is to be graphical. So let us try to summarize all these items graphically. Plot 21 shows the overall behavior for the ten counties, both alone and in contrast with all of Nebraska and all the U. S. Making this graph do triple duty has reduced its impact, but we do have our attention called to such facts as:

- the ten-county mean is usually more Republican than Nebraska as a whole.

- this difference tends to decrease in more Democratic elections.

- both Nebraska and the ten counties, were more strongly Democratic in 1932 than the whole U. S.

- Nebraska seems to tend to differ more from the total U. S. as the years roll by.

Rather than waste paper by a second copy, let us suppose that plot 17 is now repeated here, as part of our summary, completing the election-to-election results.

Turning to the county-by-county results, map 22 shows all four sets of results. (Map 20 places the counties in Nebraska, map 16 shows their names.) The geographic pattern

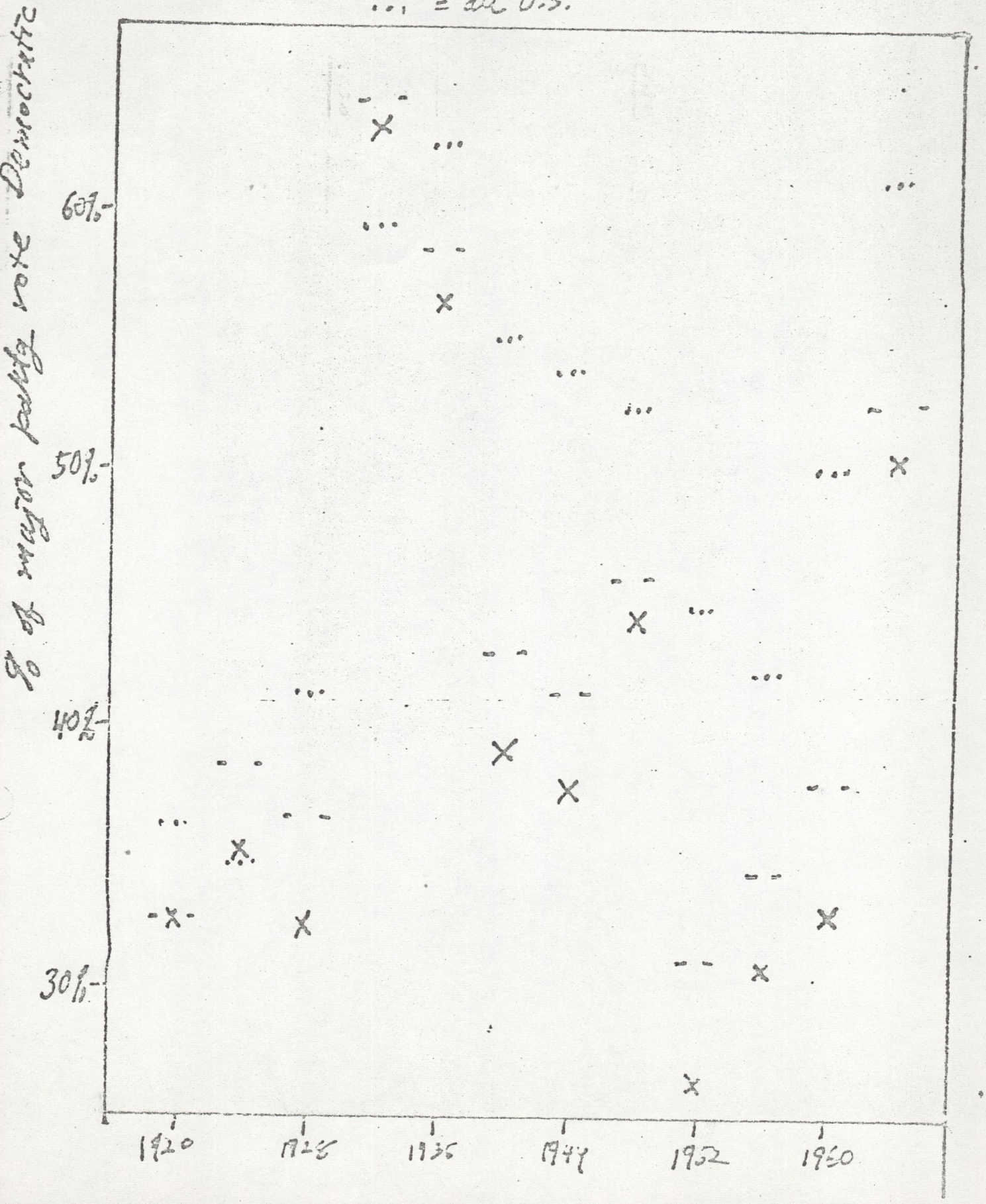
plot 21 of chapter 8

Behavior of the ten counties through time

x = ten counties

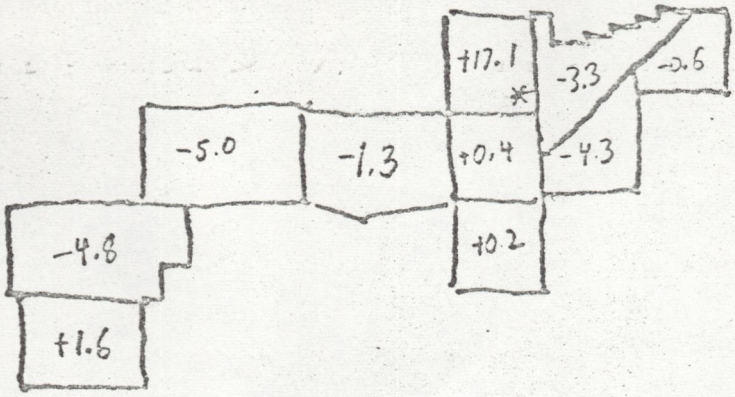
- - = all Nebraska

... = all U.S.

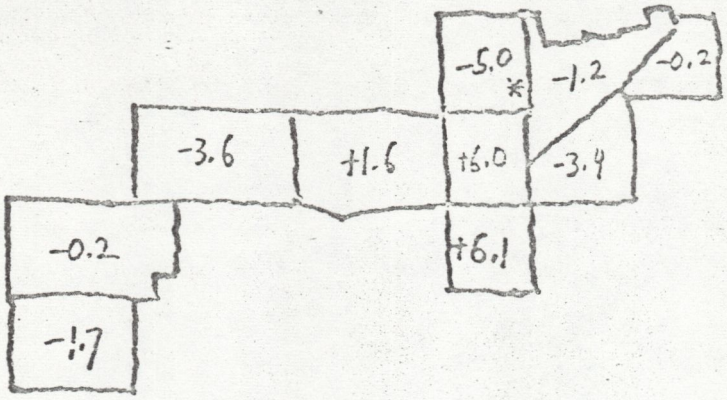


map 22 of chapter 8

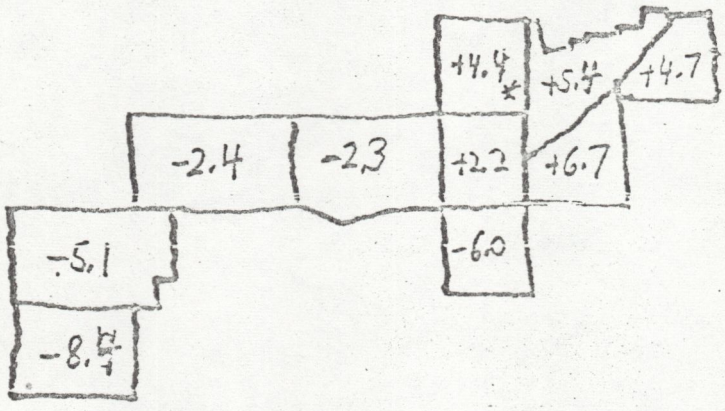
County to county comparisons



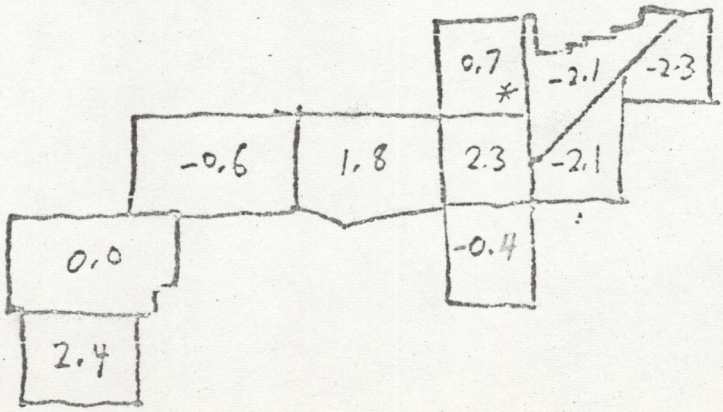
Mean residual
(+ = toward Dem)



Time Trend in
Second residuals
(+ = toward Dem)



1964 compared with
1936 in terms of
detrended residuals
(+=later more Dem)



Quiet elections (1940-1952)
compared to
quiet elections (1920-1936)
in terms of detrended
residuals.
(+ = later more Dem)

* Indicates Howard county

of county mean residuals is quite irregular, and deserves more detailed explanation than is likely to be possible without quite detailed knowledge of county characteristics. The time trends are positive in a central group of counties and negative elsewhere. The 1964-1936 comparisons are + to the northeast and - to the southwest. The "quiet election" comparisons are - to the east and + in the center and west. All three of the latter are regular enough to call for simple explanations, applying to all our 10 counties.

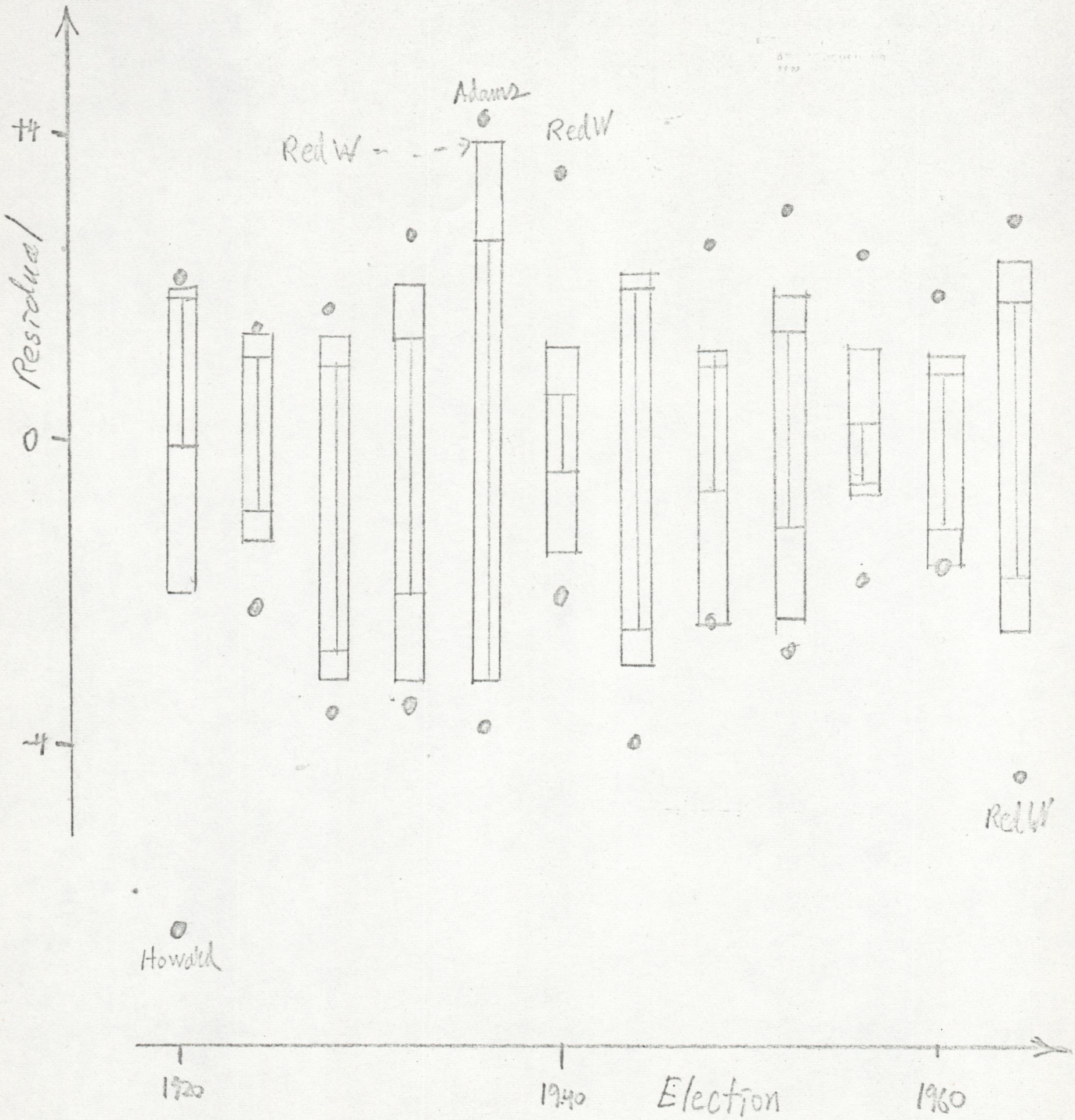
The relation of time trends to vote growth was set out in plot 12, which we may suppose repeated here.

There remains the exceptional behavior of Howard county in 1920. How exceptional was it? We may try to answer either arithmetically or graphically. Arithmetically, setting aside the 1936 and 1964 elections as unusually disturbed, only one other detrended residual is as large as ±4.0. Howard in 1920 reaches -6.5, and is thus clearly either a few or several % out of line.

Graphically, we need only turn to the technique shown in section 3B for examining spreads, and apply it to the stripes of table 13, thus displaying the residuals. (Recall that this technique plotted extreme values individually.) Plots 23 and 24 show the results of such an approach applied first to the 12 elections and then to the 10 counties. The clearly exceptional character of Howard county in 1920 is shown in both plots, as is the noticeable behavior of Red Willow county in 1964. The behavior of Adams and Red Willow in 1936 is less clearly distinctive.

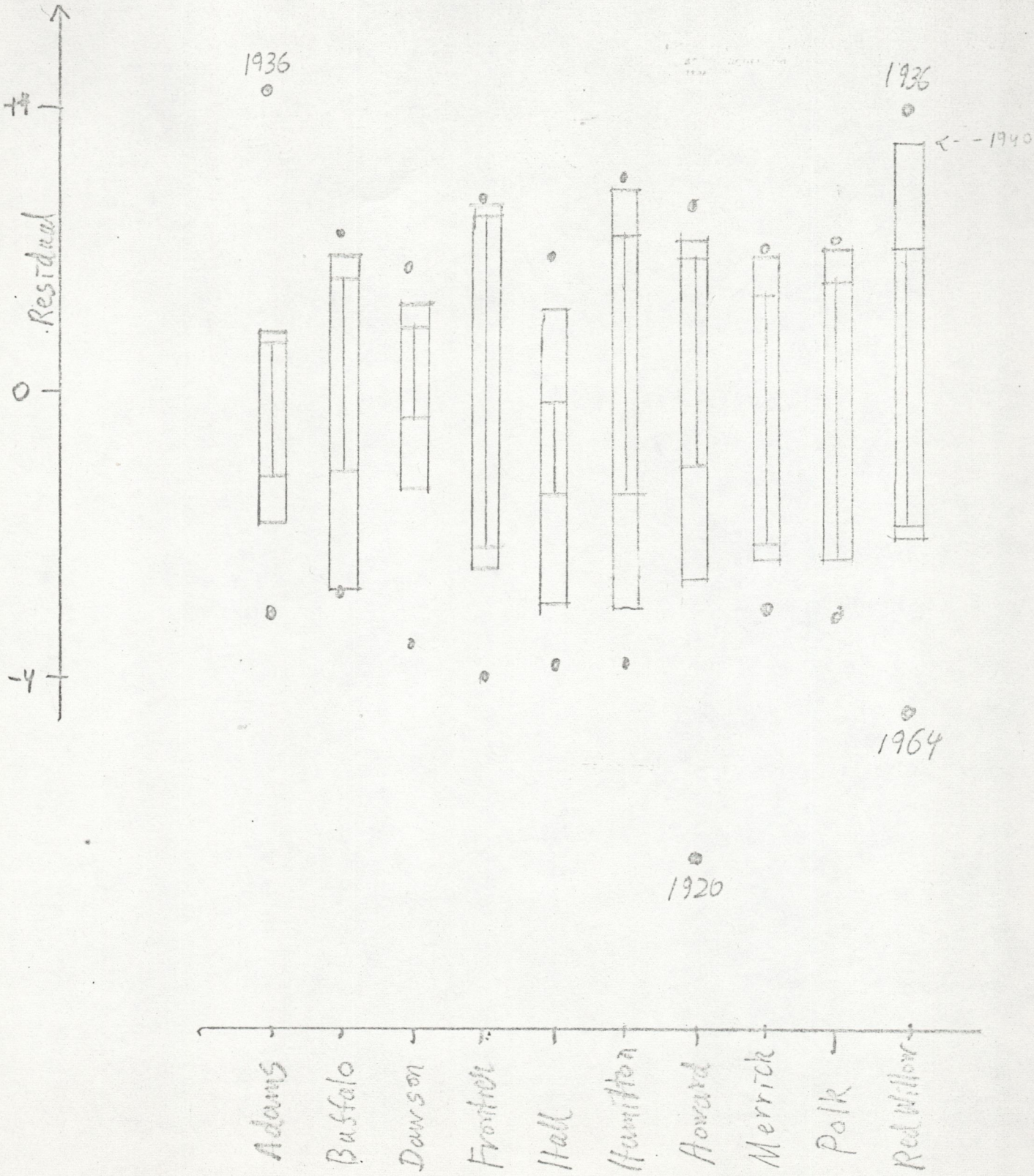
plot 23 of chapter 8

Spreads of third residuals from table 13 (election by election)



plot 24 of chapter 8

Spreads of third residuals from table 13 (county by county)



The exceptional character of Howard county in 1920 is clear. But why did this happen?

8G. Was the example gerrymandered?

The famous election district shaped like a salamander and attributed (unfairly, it is now believed) to Elbridge Gerry produced the word "gerrymander". This word was used to emphasize the care with which that district was shaped to a political end. The group of our ten contiguous counties, as our maps -- particularly map 20 -- show, is shaped more like a poodle than like a salamander, but there is rightful ground for suspicion that they were picked for a particular purpose.

Those who have not yet learned from experience how often one can peel off layer after layer of appearances and cues may think that our 10 counties were selected so that there would be layer after layer. This is not so. Rather these 10 were selected to be unusually well behaved and consistent, so much so as to allow effective analysis by means. Had we taken a different set, we might well have needed analysis by midmean differences, which would have meant rather a lot of computation for hand analysis of a 10x12 table, but we would still have found layer after layer of appearances. Selection was to keep the analysis simple, not to change its character.

8H. How far have we come?

Six sections of this chapter, and 25 tables or plots have been devoted to one example of a two-way table of responses. That there might be this much in one 10x12 table may have

surprised you. However, this sort of situation, in which layer after layer can be stripped off, is more nearly the rule than the exception. Once we acquire the structure of a two-way table, there are many things we can do -- and it is often important to do several of them in a particular instance.

We have seen examples of a number of specific techniques:

- of the peeling away of row and column means or mean residuals to reach one form of elementary analysis.
- of the use of coded tables of residuals to show to eye and brain what tables of numbers can tend to conceal.
- of the recognition and removal of trends as well as means.
- of how something recognized along one way (in response to one factor) offers us a way to ask more sharply about the apparent effect along the other way (in response to the other factor).
- of how outside information, such as the geographic arrangement of counties, can contribute to how believable (how clearly not accidental) our results are to be.

The general message is clear: Once the data has more than the simplest structure, there is much to be gained from successive layers of "clearing the trail" by peeling off one piece of partial description after another. Two-way data is much more challenging than one-way (or no-way) data and is likely to be more rewarding.